



JP2001067187

Biblio Page 1

esp@cenet

**STORAGE SUB-SYSTEM AND ITS CONTROL METHOD**

Patent Number: JP2001067187
Publication date: 2001-03-16
Inventor(s): ARAKAWA TAKASHI; MOGI KAZUHIKO; YAMAKAMI KENJI; ARAI HIROHARU
Applicant(s):: HITACHI LTD
Requested Patent: ☐ JP2001067187 (JP01067187)
Application Number: JP19990242713 19990830
Priority Number(s):
IPC Classification: G06F3/06 ; G06F12/00
EC Classification:
Equivalents:

Abstract

PROBLEM TO BE SOLVED: To simplify a work for optimizing arrangement by re-arrangement by the user of a disk array system or the like by changing the correspondence of a logical storage area from a physical storage area into the second physical storage area and executing re-arrangement.

SOLUTION: A control part 300 automatically executes re-arrangement execution processing at the set time and date. That is, the part 300 copies contents stored in a re-arrangement source physical area in a re-arrangement destination physical area based on re-arrangement information 408. Moreover, at the point of time when the copying is completed and the whole contents of the re-arrangement source physical area are reflected in the re-arrangement destination physical area, the control part 300 changes a physical area corresponding to a logical area for executing re-arrangement in logical/physical correspondence information 400 from the re-arrangement source physical area into the re-arrangement destination physical area. Besides, the control part 300 uses the re-arrangement destination physical area on a non-usage physical area 1470, changes the re-arrangement source physical area into the non-usage one and, moreover, updates the time and date of re-arrangement execution time information 406 into the one for a next time by referring to time and date updating information on re-arrangement execution time information 406.

Data supplied from the esp@cenet database - I2

16/10/01 x 2500

特開2001-67187
(P2001-67187A)

(43)公開日 平成13年3月18日(2001.3.18)

(51)Int.Cl.	識別記号	PI	フットノット(参考)
G 06 F 3/06	3 01	G 06 F 3/06	3 01 A 5 B 06 5
	5 40		5 40 5 B 08 2
	5 01	12/00	5 01 B

審査請求 未請求 請求項の数 10 O L (全 24 頁)

(21)出願番号	特開平11-242713	(71)出願人	000005108
(22)出願日	平成11年8月30日(1999.8.30)	株式会社日立製作所	株式会社日立製作所
		東京都千代田区神田駿河台四丁目6番地	東京都千代田区神田駿河台四丁目6番地
		荒川 敬史	荒川 敬史
		神奈川県横浜市東区王禅寺1095番地 株	神奈川県横浜市東区王禅寺1095番地 株
		式会社日立製作所システム開発研究所内	式会社日立製作所システム開発研究所内
		彦木 和彦	彦木 和彦
		神奈川県横浜市東区王禅寺1095番地 株	神奈川県横浜市東区王禅寺1095番地 株
		式会社日立製作所システム開発研究所内	式会社日立製作所システム開発研究所内
		(74)代理人 100075096	(74)代理人 100075096
		弁理士 作田 康夫	弁理士 作田 康夫

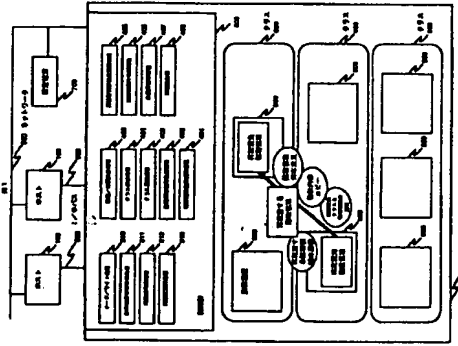
最終頁に続く

(54)【発明の名称】 ストレージサブシステム及びその制御方法

(57)【要約】

【課題】 ストレージサブシステムのユーザまたは保守員が記憶領域の物理的再配置による配置最適化を行うための作業を簡便にするストレージサブシステムおよび制御方法を提供する。

【解決手段】 ストレージサブシステム200は、記憶装置500を、それぞれ属性を有する複数の組(クラス)600として管理し、クラス属性に基づき最適な再配置先のクラスを決定する。



【特許請求の範囲】

【請求項1】複数の記憶装置と、前記記憶装置の使用状況情報を取得する手段と、前記計算機がリードライต์対象とする記憶装置と前記記憶装置の第一の物理記憶領域との対応づけを行う手段とを有し、1台以上の計算機に接続するストレージサブシステムの制御方法であって、前記記憶装置は複数の組(クラス)に分類され、前記クラスは設定された属性を有し、前記ストレージサブシステムは、前記使用状況情報および前記クラス属性に基づき前記記憶装置の再配置先として利用可能な第一の物理記憶領域の内容を前記第二の物理記憶領域にコピーするとともに論理記憶領域の対応づけを前記第一の物理記憶領域から前記第二の物理記憶領域へ変更して再配置を行うことを特徴とするストレージサブシステムの制御方法。

【請求項2】請求項1に記載のストレージサブシステムは、前記制御方法であって、ストレージサブシステムは、前記使用状況情報を蓄積し、設定された期間の前記使用状況情報に基づき、論理記憶領域の再配置先を決定し、設定された時間に再配置を行うことを特徴とするストレージサブシステムの制御方法。

【請求項3】請求項1または2に記載のストレージサブシステムは、使用状況情報として、記憶装置の単位時間当たりの使用時間(使用率)を用い、各クラスは、属性として設定されたクラス間の性能順位と使用率上限値を有し、前記ストレージサブシステムは、クラスの使用率上限値を超えている記憶装置から再配置する論理記憶領域を選択し、前記論理記憶領域の再配置先のクラスを前記順位の高いクラスから、各クラスの使用率上限値を超えないように決定することを特徴とするストレージサブシステムの制御方法。

【請求項4】請求項1または2に記載のストレージサブシステムの制御方法であって、ストレージサブシステムは、使用状況情報として、記憶装置の単位時間当たりの使用時間(使用率)を用い、各クラスは、属性として設定されたクラス間の性能順位と使用率上限値を有し、前記ストレージサブシステムは、クラスの使用率上限値を超えている記憶装置から再配置する論理記憶領域を選択し、前記論理記憶領域の再配置先として利用可能な物理記憶領域を同一クラス内の記憶装置から、前記クラスの使用率上限値を超えないように決定することを特徴とするストレージサブシステムの制御方法。

【請求項5】請求項1または2に記載のストレージサブシステムの制御方法であって、記憶装置の単位時間当たりの使用時間(使用率)を用い、各クラスは属性として設定された対象アクセス種別と使用率上限値を有し、前記ス

トレージサブシステムは、クラスの使用率上限値を超えている記憶装置から再配置する論理記憶領域を選択し、前記論理記憶領域に対するアクセス種別の分析結果に基づいて前記記憶装置の再配置先のクラスを前記対象アクセス種別のクラスから、各クラスの使用率上限値を超えないように決定することを特徴とするストレージサブシステムの制御方法。

【請求項6】1台以上の計算機に接続し、複数の記憶装置と、前記記憶装置の使用状況情報を取得する手段と、前記計算機がリードライต์対象とする論理記憶領域と、前記記憶装置の第一の物理記憶領域との対応づけを行う手段とを有するストレージサブシステムであって、前記複数のディスク装置をそれぞれ属性を有する複数の組(クラス)として管理する手段と、前記使用状況情報および前記クラス属性に基づき前記論理記憶領域に好適な再配置先のクラスを決定する手段と、前記論理記憶領域の再配置先として利用可能な第二の物理記憶領域を前記クラス内から選択する手段と、前記第一の物理記憶領域の内容を前記第二の物理記憶領域にコピーするとともに論理記憶領域の対応づけを前記第一の物理記憶領域から前記第二の物理記憶領域へ変更して再配置を行う手段とを有することを特徴とするストレージサブシステム。

【請求項7】請求項6に記載のストレージサブシステムであって、ストレージサブシステムは、前記使用状況情報を蓄積し、設定された期間の前記使用状況情報に基づき、論理記憶領域の再配置先を自動的に決定する手段と、設定された時間に再配置を行う手段とを有することを特徴とするストレージサブシステム。

【請求項8】請求項6または7に記載のストレージサブシステムであって、ストレージサブシステムは、使用状況情報として記憶装置の単位時間当たりの使用時間(使用率)を用いる手段を有し、前記ストレージサブシステムは、各クラスに属性として設定されている使用率上限値を超えている記憶装置から再配置する論理記憶領域を選択する手段と、前記論理記憶領域の再配置先のクラスを各クラスに属性として設定されているクラス間の性能順位から、各クラスの使用率上限値を超えないように決定する手段とを有することを特徴とするストレージサブシステム。

【請求項9】請求項6または7に記載のストレージサブシステムであって、ストレージサブシステムは、使用状況情報として、記憶装置の単位時間当たりの使用時間(使用率)を用いる手段を有し、前記ストレージサブシステムは、属性として設定されたクラスの使用率上限値を超えている記憶装置から再配置する論理記憶領域を選択する手段と、対象アクセス種別を属性として設定されたクラスから、前記論理記憶領域の再配置先のクラスを前記分析結果に基づいて各クラスの使用率上限値を超えないように決定する手段とを有することを特徴とする

045'880/03

るストレージサブシステム。

【請求項10】請求項6、7、8、または9に記載のストレージサブシステムであって、ストレージサブシステムは、複数のディスク装置を有するディスクアレイであり、前記ディスク装置の使用率を使用状況情報として用いる手段を有することを特徴とするストレージサブシステム。

【発明の詳細な説明】

【0001】
【発明の属する技術分野】本発明は、複数の記憶装置を有するストレージサブシステム、およびその制御方法に関する。

$$\{0002\}$$

【従来の技術】コンピュータシステムにおいて、高性能を実現する二次記憶システムの一つにディスクアレイシステムがある。ディスクアレイシステムは、複数のディスク装置をアレイ状に配置し、前記各ディスク装置に分割格納されるデータのリード/ライトを、前記各ディスク装置を並列に動作させることによって、高速に行うシ

としては、D. A. Patterson, G. Gibbons, and R. H. Kats, "A Case for Redundant Arrays of 1

expensive Disks (RAID) *
(in Proc. ACM SIGMOD, pp.
109-116, June 1988)がある。この
論文では、冗長性を付加したディスクアレイシステムに
対し、その構成に応じてレベル1からレベル5の種別を
与えている。これらの種別に加えて、冗長性無しのディ

の最適配置を実現するディスプレイシステムが開示されている。

[000]

【発明が解決しようとする課題】特開平9-274544号公報に示されるような従来の技術における配置最適化の実行方法については以下の課題がある。

【0006】再配置する論理記憶領域の選択および再配
置先の物理記憶領域の選択にあたり、ディスクアレイシ
ステムのユーザまたは保守員が、前記ディスクアレイシ
ステムの構成や個々のディスク装置の特性や性能などの
情報を確認して前記選択を行わなければならない。ユーザ
または保守員による作業が煩雑となっていた。

【0007】また、ディスクアレイシステムが選択を自動的に行う場合においても、ユーザまたは保守員が情報交換のためのディスク装置の接続して選択基準値を規定するディスク装置の装着の情報を随時または保守員による作業が煩雑となっていた、特に、上記のように異種のレベールや異種のディスク装置の混在するディスクアレイシステムについては情報管理の煩雑さが増大する。

【0008】また、ディस्कアレイシステムが選択のために、ディスクアレイシステムの参照は、ホストコンピュータおよびディスクアレイシステムの特性を考慮してなかつた。一般にコンピュータシステムで行われる処理と処理に伴う1/Oは、ユーザによって作成されたスナジュールに則って行われており、また処理および1/Oの傾向は日毎、月毎、年毎などの周期性を示す場合も多く、一週間にユーザが特定期間の処理および1/Oに関心があると考えられる。

【0009】また上記従来技術において、再配置による物理的再配置方法については以下の課題がある。物理的再配置による性能チューニング方法は、ディスク装換、すなわち、物理記憶領域の使用状況に変更を加えるものであるが、従来の技術においては、ホストコンピュータから物理記憶領域に對する1/Oアクセスについての情報も参照する物理、再配置する論理記憶領域の選定および再配置優先の物理、再配置する論理記憶領域の選定が与えられない可能性があった。

【0010】 また、ホストコンピュータからのシグナルをシャトルアクセスとランダムアクセスが顕著に、同一のデイスクリューズ装置に含まれる別々の物理配憶領域に対して行われる場合でも、シグナルアクセスとランダムアクセスと先ず異なるディस्क装置に分離するために、再度優先せしめる必要のあるディスク装置に特定して自動的再配置を行わせることはできなかった。一般に、ホストコンピュータから送られる処理要件として、データ長の小さいランダムアクセスよりも長時間での応答（高応答性）が求められるが、ランダムアクセス装置にデータ長の大きいシーケンシャルアクセスが存在する場合、ランダムアクセスの応答時間はシーケンシャルアクセスの処理に阻害されて長くなり、

応答性能は悪化してしまう。

【0011】本発明の第一の目的は、ディスクアレイシステムのエューザーまたは保守員が再配置による配置最適化を行うための作業を簡便することにある。

【0012】本発明の第二の目的は、ホストコンピュータおよびディスクレイシステムを含むシステムでの処理のスケジュールを考慮した再配置による配置最適化を可能にすることにある。

【0013】本発明の第三の目的は、再配置する論理記憶領域の選択および再配置先の物理記憶領域の選択にある。実際の記憶装置であるディスクレイシスシステムの制御方法に基づき選択を行う。ディスクレイシスシステムを提供することにある。

【0014】本発明の第四の目的は、ディスクレイシスシステムにおける同一ディスク装置での顕著なシーケンシャルアクセスとランダムアクセスの存在に対し、再配置先のディスク装置を任意に特定して再配置によりシーケンシャルアクセスおよびランダムアクセスを異なるディスク装置に自動的に分離することができるようにすることにある。

【0015】
 【課題を解決するための手段】上記の第一目的を実現す
 るために、1台以上のホストコンピュータに接続するデ
 ィスクデバイスレボリューションは、配下のホストコンピュ
 ータが使用状況情報を受得する手段と、ホストコンピュータ
 リード/ライト対象とする論理記憶領域とディスタ装置
 の第一の物理記憶領域との対応づけを行う手段とを有
 し、さらに、複数のディスク装置をそれぞれ属性を有す
 る複数の組（クラス）として管理する手段と、使用状況
 情報およびクラス属性に基づき論理記憶領域に好適な再
 配置先のクラスを決定する手段と、論理記憶領域の再配
 置先として利用可能な第二の物理記憶領域をクラスか
 ら選択する手段と、第一の物理記憶領域の内容を前記第
 二の物理記憶領域にコピーする手段とにも論理記憶領
 域の対応づけを第一の物理記憶領域から第二の物理記憶
 領域へ変更して再配置を行う手段を備える。

【0016】また、上記第二の目的を實現するのために、ディスクアレイシステムは、使用状況情報を蓄積し、設定された期間の使用状況情報に基づき、論理記憶領域の再配置先を決定する手段と、設定された時間に再配置を行う手段を備えることができる。

【0017】また、上記第三の目的を実現するために、ディस्कアクセスシステムは、使用状況情報として、ディस्क装置の単位時間当たりの使用時間（使用率）を用いる手段を備える。

【0018】また、上記第四の目的を実現するために、ディスタンスシステムは、各クラスに属性として設定された対象アクセス種別（シーケンシャル/ランダムアクセス種別）と使用率上限値を用いて、クラスの使用率上限値を超えている記憶装置から再配属する記憶装置

域を選択し、論理記憶領域に対するアクセス種別を分析結果に基づいて論理記憶領域の再配置先のクラスを好適なアクセス種別のクラスから、各クラスの使用率上限値を超えないように決定する手段を備える。

[0019]

【発明の実施の形態】以下、本発明の実施の形態を図1～図27を用いて説明する。

【0020】＜第一の実施の形態＞本実施の形態では、クラス600に基づく再配置の判断と、再配置判断および実行のスケジューリングについて説明する。

【0021】図1は、本発明の第1の実施の形態における計算機システムの構成図である。

【0022】本実施の形態における計算機システムは、ホスト100、ストレージサブシステム200、制御端末700を有してなる。

【0023】ホスト100は、ストレージサブシステム200に1/〇バス800を介して接続し、ストレージサブシステム200に対してリードやライトの1/〇を行う。1/〇の際、ホスト100は、ストレージサブシステム200の記憶領域について論理領域を指定する。1/〇バス800の例としては、ESCON、SCSI、ファイバチャネルなどがある。

【0024】ストレージサブシステム200は、制御部3300および複数の記憶装置500を有する。制御部3

310は、リード/ライト処理310、使用状況情報取得処理311、再配置判断処理312、及び再配置実行処理313を含む。

理313を行う。また、ストレーザブシテム200は、論理／物理対応情報400、クラス構成情報40

1、クラス属性情報 402、論理領域使用状況情報 403、物理領域使用状況情報 404、再配置判断対象期間情報 405、再配置実行時刻情報 406、未使用領域情報 407、及び再配置情報 408を保持する。

【0025】ホスト100、制御部300、および制御部700は、ネットワーク900で接続される。ネットワーク900の例としては、FDDI、ファイバチャネルなどがある。

いても、性能や容量などによりコストが異なるため、ディスクアクセスシステム201を構成するにあたって最適なコストパフォーマンスを実現するために性能や容量の異なる複数のディスク装置502を用いることもあ
る。よって本実施の形態においてディスクアクセスシステム201を構築する各パーティグループ501は性能、信頼性、特性などの属性が同一であるとは限らず、特に性能について差異があるとする。

【0050】本実施の形態における論理/物理対応情報400の一例を図11に示す。

【0051】論理アドレスは、ホスト100がリード/ライト処理310で用いる論理領域を示すアドレスである。物理アドレスは実際にデータと前記冗長データが格納されるディスク装置502上の領域を示すアドレスであり、パーティグループ番号と各々のディスク装置番号およびディスク装置内アドレスからなる。パーティグループ番号は個々のパーティグループ501を示す。ディスク装置番号は個々のディスク装置502を示す。ディスク装置内アドレスはディスク装置502内の領域を示すアドレスである。制御部300は、RAIDの動作として、冗長データに関する情報を前記リード/ライト処理310などで用いて処理するが、本実施の形態の図明では、パーティグループ502を動作上の1単位として説明する。前記処理に関してはここでは特に示さない。

【0052】さらに第1の実施の形態と同様に、パーティグループ501は、ユーザによってまたは初期状態として複数の組(クラス600)に分類されておられ、クラス600への分類はクラス構成情報401に設定されている。クラス構成情報401に設定されている。【0053】クラス番号は各クラス600を示す番号である。パーティグループ数は各クラス600に属するパーティグループの数を示す。パーティグループ番号は各クラス600に属するパーティグループ番号501を示す。同様に、各クラス600の属性は、クラス属性情報402に設定されている。本実施の形態におけるクラス属性情報402の一例を図13に示す。

【0054】クラス番号は、各クラス600を示す番号である。使用上限値は後述のディスク使用率の許容範囲を示す上限値であり、クラス600の属するパーティグループ501に適用する。クラス間性能順位は、クラス600間の性能順位(数字の小さいものが高性能とす)である。クラス間性能順位は各クラス600を構成するパーティグループ501の前述の性能属性に基づき、所配実行上限値および固定については後述する。【0055】本実施の形態における使用状況取得情報311について図14で説明する。

【0056】制御部300は、第1の実施の形態と同様に、リード/ライト処理310において使用したディスク装置502の使用時間取得して単位時間当たりの使

用時間(使用率)を求め、さらに、ディスク装置502が属するパーティグループ501について、使用率の平均を算出し(ステップ1300)、使用率平均を、リード/ライト対象となった論理領域についてのディスク装置使用率として論理領域使用状況情報403に記録する(ステップ1310)。また制御部300は、パーティグループ501に対応する全論理領域のディスク装置使用率の和を求め(ステップ1320)、パーティグループ501の使用率として物理領域使用状況情報404に記録する(ステップ1330)。

【0057】本実施の形態における論理領域使用状況情報403および物理領域使用状況情報404の一例を図15および図16に示す。

【0058】日時とはサンプリング間隔(一定期間)毎の日時を示し、論理アドレスは論理領域を示し、パーティグループ番号は個々のパーティグループを示し、論理領域のディスク装置使用率およびパーティグループ使用率はそれぞれ前記サンプリング間隔での平均使用率を示す。上記のようなディスク装置502の使用率は、ディスク装置502にかかると負荷を示す値であり、使用率が大きい場合は、ディスク装置502が性能ボトルネックとなっている可能性があるため、再配置処理で使用率を下げることによりディスクアクセスシステム201の性能向上が期待できる。

【0059】次に、再配置判断処理312について図17で説明する。

【0060】制御部300は、各クラス600について、クラス600に属するパーティグループ501をクラス構成情報401から取得する(ステップ1300)。続いて、制御部300は、第1の実施の形態と同様の再配置判断対象期間情報405を参照して対象期間を取得し、さらにパーティグループ501について、対象期間の物理領域使用状況情報404のパーティグループ使用率を取得し集計する(ステップ1320)。続いて、制御部300は、クラス属性情報402を参照してクラス600の使用率上限値を取得する(ステップ1330)。制御部300は、パーティグループ使用率とクラス上限値と比較し、パーティグループ使用率がクラス上限値より大きい場合は、パーティグループ501の使用率を減らすために、パーティグループ501に対応する論理領域の再配置が必要と判断する(ステップ1340)。

【0061】続いて、制御部300は、対象期間の論理領域使用状況情報403を参照して、再配置が必要と判断したパーティグループ501の各物理領域に対応する論理領域のディスク装置使用率を取得し集計し(ステップ1350)、ディスク装置使用率の大きいもの(スケ

ラ、再配置する論理領域として選択する(ステップ1360)。論理領域の選択は、パーティグループ501の使用率から選択した論理領域のディスク使用率を減算し

ていき、クラス600の使用率上限値以下になるまで行う(1370)。ディスク装置使用率の大きい論理領域は、パーティグループ501の使用率に対する影響も大きく、またホスト100からの論理領域に対するアクセス頻度も大きいと考えられるため、ディスク装置使用率の大きい論理領域を優先的に再配置することで、ディスクアクセスシステム201の効果的な性能改善が期待できる。

【0062】制御部300は、選択された論理領域についての物理領域となる物理領域を探し、パーティグループ501の属するクラス600より性能順位が高位のクラス600(高性能クラス)に注目し、クラス構成情報401および第1の実施の形態と同様の未使用物理領域情報407を参照して高性能クラスに属するパーティグループ501の未使用物理領域を取得する(ステップ1380)。

【0063】さらに、制御部300は、各未使用物理領域について、再配置先とした場合のパーティグループ使用率の平均値を求め(ステップ1390)、未使用物理領域の中から、再配置先とした場合に高性能クラスに設定されている上限値を超えないと予測できる未使用物理領域を、再配置先物理領域として選択し(ステップ1400)、選択結果を第1の実施の形態と同様に、再配置情報408に出力する(ステップ1410)。選択した全ての論理領域について再配置先の物理領域を選択し終えたら処理を終了する(ステップ1420)。

【0064】本実施の形態において、制御部300は、第1の実施の形態に加えてパーティグループ情報409を保持し、パーティグループ情報409、論理領域使用状況情報403、及び物理領域使用状況情報404から使用率予測値を算出する。

【0065】パーティグループ情報409の一例を図18に示す。パーティグループ番号は個々のパーティグループ501を示す番号である。RAID構成はパーティグループ501が構成するRAIDのレベルやディスク台数や冗長度構成を示す。ディスク装置性能はパーティグループ501を構成するディスク装置502の性能特性を示す。固定については後述する。上記の処理においてディスク装置使用率の大きい論理領域の再配置先を高性能クラスのパーティグループ501とすることで、同一負荷に対するディスク装置使用時間を短縮でき、論理領域の再配置後のディスク装置使用率を抑制できる。

【0066】再配置実行処理313は、第1の実施の形態と同様に行われるが、図19に示すように、制御部300は、再配置のためのコピーを行う前にクラス属性情報402を参照し、再配置元および再配置先のクラス600について、ユーザによってまたは初期条件として設定された再配置実行上限値を取得する(ステップ1500)。さらに物理領域使用状況情報404を参照して、

再配置元および再配置先のパーティグループ501の直近のパーティグループ使用率を取得し(ステップ1510)、比較の結果少なくとも一方のクラス600においてパーティグループ使用率が再配置実行上限値を超えていた場合は(ステップ1520、1530)、再配置実行処理313を中止または延期する(ステップ1540)。

【0067】上記処理によりユーザは、パーティグループ501の使用率が大きくすくなら負荷が高い場合に向記コピーによりさらに負荷が生じることを回避することができ、また回避のための上限値をクラス600毎に任意に設定することができる。

【0068】上記のように処理することによって、ディスク装置502の使用状況に基づいて物理的に再配置する論理領域の選択、および再配置先の物理領域の選択を、クラス構成および属性に基づいて行い、再配置によりディスク装置502の負荷を分散して、各クラス600に設定されている使用率上限値を、クラス600に属するパーティグループ501の使用率が超えない配置を実現することができ、さらに再配置判断および実行の処理を繰り返して配置を修正していくことによって、使用状況の変動や予測誤差を吸収していくことができる。

【0069】再配置判断処理312において、制御部300は、対象期間の物理領域使用状況情報404のパーティグループ使用率や、論理領域使用状況情報403の論理領域のディスク装置使用率を参照して集計し、判断に用いるとしたが、例えば、対象期間の全ての値の平均を用いる代わりに、対象期間中の上位m割の値を用いる方法も考えられ、また上位m番目の値を用いる方法も考えられる(mは1以上の整数)。これらの方法をユーザが選択できるようにすることで、ユーザは使用状況の特性的な部分のみを選択して用い、再配置判断処理312を行わせることができる。

【0070】上記の再配置判断処理312において、制御部300は、ディスクアクセスシステム201の全てのクラス600について、論理領域の再配置の必要なパーティグループ501の検出を行うとしたが、前記検出の前に制御部300がクラス属性情報402を参照し、固定属性が設定されているクラス600については、検出の対象外としてもよい。また同様に、制御部300がパーティグループ情報409を参照し、固定属性が設定されているパーティグループ501については検出の対象外としてもよい。また、再配置判断処理313において、制御部300は、高性能クラスに属するパーティグループ501の未使用物理領域から再配置先の物理領域を選択としたが、固定属性が設定されているクラス600については対象外として、さらに性能順位が高位のクラス600を高性能クラスとして扱い、パーティグループ501については対象外としてもよい。上記のように固

定属性が設定されているクラス600またはパリティグループ501を扱うことにより、ユーザは上記の自動的な再配置処理において物理的な再配置の影響を生じさせたくないクラス600またはパリティグループ501を指定し、再配置の対象外とすることができる。

【0071】＜第3の実施の形態＞本実施の形態では、同一クラス600内での再配置判断について説明する。本実施の形態での計算機システムは、第2の実施の形態と同様である。ただし、本実施の形態では1つのクラス600に複数のパリティグループ501が属する。本実施の形態での処理は、再配置判断処理312を除いては第2の実施の形態と同様である。また、再配置判断処理312についても、再配置する論理領域の選択（ステップ1600）は、第2の実施の形態と同様である。

【0072】本実施の形態での再配置判断処理312における、再配置先の物理領域の選択について図20で説明する。

【0073】第2の実施の形態では再配置先の物理領域を再配置元の物理領域の属するクラス600より性能値が低い位のクラス600から選択するが、本実施の形態では同一クラス600の再配置元以外のパリティグループ501から選択する。制御部300は、クラス構成情報401と未使用領域情報407を参照して、同一クラス600に属する再配置元以外のパリティグループ501の未使用物理領域を取得する（ステップ1610）。

制御部300は、各未使用物理領域について、再配置先とした場合のパリティグループ501の使用率の測定値を求め（ステップ1620）、未使用物理領域の中から、再配置先とした場合に同一クラス600に設定されている上限値を超えないと予測できる未使用物理領域を、再配置先の物理領域として選択し（ステップ1630）、選択結果を第2の実施の形態と同様に、再配置情報408に出力する（ステップ1640）。再配置する全ての論理領域について再配置先の物理領域を選択し終えたら処理を終了する（ステップ1650）。

【0074】上記の処理により、同一クラス600内においてディスク装置502の負荷を分散することができる。上記の処理方法は例えばディスクアレイシステム201のパリティグループ501が全て1つのクラス600（単一グループ）に属する構成に適用することができる。また、例えば、第2の実施の形態で説明した処理方法と組み合わせた場合に、再配置先の未使用物理領域の選択において、再配置元のクラス600より性能値が低い位のクラス600に最適な未使用物理領域が得られない場合や、性能値が最上位のクラス600での処理に適用できる。第2の実施の形態で説明した処理方法と組み合わせた場合は、第2の実施の形態での処理方法と異なる使用率上限値を用いてもよく、すなわち、そのためにクラス属性情報402が各クラス600について

二種類の使用率上限値または差分を有してもよい。

【0075】＜第4の実施の形態＞本実施の形態では、第2の実施の形態での再配置判断処理312において、再配置元のクラス600より性能値が低い位のクラス600（高性能クラス）に再配置先の未使用物理領域が見つからなかつた場合に、再配置先を得るために先立って行われる、性能値がより低い位のクラス600（低性能クラス）への高性能クラスからの再配置の処理について説明する。

【0076】本実施の形態での計算機システムは、第2の実施の形態と同様である。本実施の形態における再配置判断処理312について図21で説明する。

【0077】制御部300は、高性能クラスに属するパリティグループ501をクラス構成情報401から取得する（ステップ1700）。続いて制御部300は、第1の実施の形態と同様の再配置判断対象期間情報405を参照して対象期間を取得し（ステップ1710）、対象期間の論理領域の使用状況情報403を参照して、パリティグループ501の各物理領域に対応する論理領域のディスク装置使用率を取得し（ステップ1720）、ディスク装置使用率の小さいものから、低性能クラスへ再配置する論理領域として選択する（ステップ1730）。このとき論理領域の選択は必要だけ行われる（ステップ1740）。

【0078】続いて制御部300は、選択された論理領域についての再配置先となる物理領域を、低性能クラスに属するパリティグループ501から選択するが、再配置先の物理領域選択の処理は、第2の実施の形態での処理説明において再配置先としている高性能クラスを低性能クラスと読み替えれば、第2の実施の形態での処理と同様である（ステップ1750）。また、本実施の形態におけるその他の処理も第2の実施の形態での処理と同様である。

【0079】上記の処理を行うことで、第2の実施の形態での再配置判断処理312において高性能クラスに再配置先の未使用物理領域が見つからなかつた場合に、高性能クラスから低性能クラスへ論理領域の再配置を、高性能クラスへの再配置に先立って行い、再配置先の未使用物理領域を高性能クラスに用意することができる。制御部300は、上記の処理を必要に応じて繰り返して、十分な未使用物理領域を用意することができる。

【0080】論理領域の再配置先を低性能クラスのパリティグループ501とするため、同一負荷に対するディスク使用時間が再配置について増大し、論理領域の再配置後のディスク装置使用率が增大する可能性があるが、ディスク使用率の小さい論理領域から再配置していくことで、増大の影響を最小限に抑えることができる。

【0081】＜第5の実施の形態＞本実施の形態では、クラス600の属性の1つにアクセス種別属性を設け、

アクセス種別属性を用いてシーケンシャルアクセスが顕著に行われる論理領域とランダムアクセスが顕著に行われる論理領域とを、他のパリティグループ501に自動的に物理的に再配置して分離するための再配置判断について説明する。

【0082】本実施の形態における計算機システムは図10に示したものである。本実施の形態では、第2の実施の形態での説明に加え、制御部300が保持する下記の情報を用いる。

【0083】本実施の形態でのクラス属性情報402の一例を図22に示す。この例では、第2の実施の形態の例に対しアクセス種別が加えられており、クラス600のアクセス種別が、例えばシーケンシャルアクセスである場合は、クラス600がシーケンシャルアクセスに好適であると設定されていることを示す。

【0084】本実施の形態での論理領域使用状況情報403の一例を図23に示す。この例では、第2の実施の形態の例に対し、シーケンシャルアクセス率およびランダムアクセス率が加えられている。

【0085】さらに、本実施の形態において制御部300は、第2の実施の形態に加え、アクセス種別基準値情報410と論理領域属性情報411を保持する。

【0086】アクセス種別基準値情報410の一例を図24に示す。ユーザによりまたは初期条件として、アクセス種別基準値情報410には後述のアクセス種別の判定に用いる基準値が設定されている。また、論理領域属性情報411の一例を図25に示す。アクセス種別ヒントは、各論理領域について顕著に行われると期待できるアクセス種別であり、ユーザが設定する。固定については後述する。

【0087】本実施の形態での処理は、使用状況情報取得処理311および再配置判断処理312を除いては第2の実施の形態と同様である。

【0088】本実施の形態における使用状況情報取得処理311について図26で説明する。

【0089】制御部300は、第2の実施の形態での使用状況情報取得処理311と同様に、論理領域についてのディスク装置使用率を算出し（ステップ1800、1810）、リード/ライト処理310の使用率内容を分析して、使用率についてシーケンシャルアクセスとランダムアクセスの比率を算出し（ステップ1820）、使用率およびアクセス種別比率を論理領域使用状況情報403に記録する（ステップ1830）。また、制御部300は、第2の実施の形態と同様にパリティグループ501の再配置先と物理領域使用状況情報404への記録を行う（ステップ1840、1850）。

【0090】本実施の形態における再配置判断処理312において、再配置する論理領域の選択は第2の実施の形態と同様である（ステップ1900）。再配置判断処理312での再配置先の物理領域の選択について図27

で説明する。

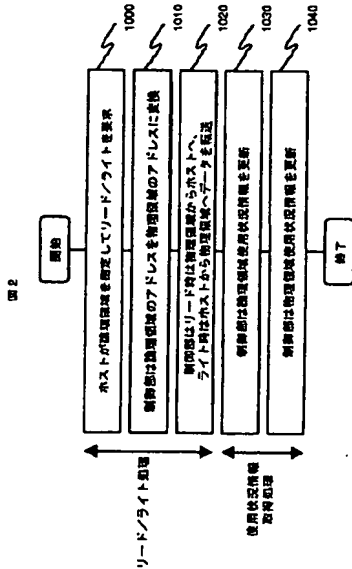
【0091】制御部300は、論理領域使用情報403を参照し、再配置する論理領域についてのシーケンシャルアクセス率を取得し（ステップ1910）、アクセス種別基準値情報410に設定されている基準値と比較する（ステップ1920）。シーケンシャルアクセス率が基準値より大きい場合、制御部300は、クラス属性情報402を参照し、アクセス種別がシーケンシャルと設定されているクラス600（シーケンシャルクラス）が存在するかどうかを（ステップ1950）。シーケンシャルクラスが存在する場合、制御部300は、クラス構成情報401と未使用物理領域情報407を参照して、シーケンシャルクラスに属する再配置元以外のパリティグループ501の未使用物理領域を取得する（ステップ1960）。さらに制御部300は、各未使用物理領域について、再配置先とした場合のパリティグループ使用率の測定値を求め（ステップ1970）、未使用物理領域の中から、再配置先とした場合にシーケンシャルクラスに設定されている上限値を超えないと予測できる未使用物理領域を、再配置先の物理領域として選択し（ステップ1980）、選択結果を第2の実施の形態と同様に再配置情報408に出力する（ステップ1990）。制御部300は、使用率測定値を、第2の実施の形態と同様のパリティグループ情報409と本実施の形態における論理領域使用状況情報403および物理領域使用状況情報404から算出する。

【0092】前記の比較において、シーケンシャルアクセス率は基準値以下である場合、制御部300は、論理領域属性情報411を参照し、論理領域についてアクセス種別ヒントがシーケンシャルと設定されているか調べ（ステップ1940）。アクセス種別ヒントにシーケンシャルと設定されている場合、上記と同様に制御部300は、シーケンシャルクラスの有無を調べ（ステップ1950）、シーケンシャルクラスが存在する場合は、シーケンシャルクラスから再配置先の物理領域を選択する（ステップ1960～1990）。

【0093】前記の比較において、シーケンシャルアクセス率が前記基準値以下であり、さらにアクセス種別ヒントがシーケンシャルでなかつた場合、またはシーケンシャルクラスが存在しなかつた場合、制御部300は、第2の実施の形態と同様に、シーケンシャルクラスの物理領域を選択する（ステップ2000）。

【0094】上記の処理により、同一パリティグループ501での顕著なシーケンシャルアクセスとランダムアクセスの混在に対し、各クラス600に属性として設定されたアクセス種別と使用率上限値を用いて、ランダムアクセスが顕著に行われる論理領域と、ランダムアクセスが顕著に行われる論理領域とを、異なるパリティグループ501に自動的に再配置して分離、すなわち異

【図2】



【図3】

図3

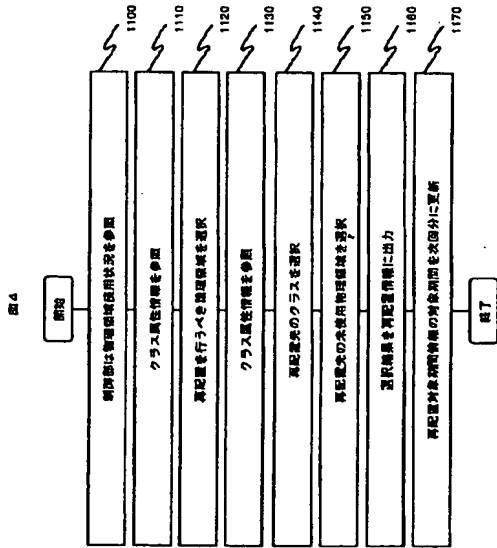
物理アドレス	物理アドレス	
	記憶装置番号	記憶装置内アドレス
0~999	0	0~999
1000~1999	0	1000~1999
2000~2999	1	0~999
3000~3999	1	1000~1999

【図5】

図5

開始日時	1999年8月11日 0時30分
終了日時	1999年8月11日 17時15分
時間更新情報	毎日(+24時間)

【図4】



【図6】

図6

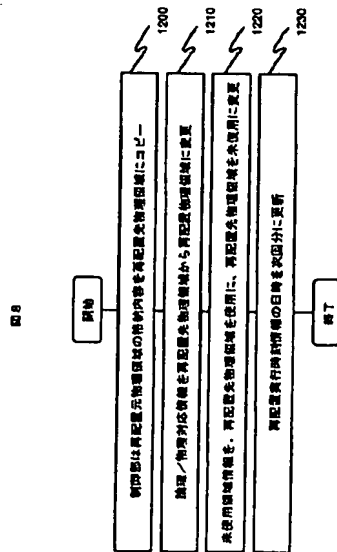
番号	記憶装置	物理装置情報		物理装置内物理装置	
		記憶装置番号	記憶装置内アドレス	記憶装置番号	記憶装置内アドレス
1	0~999	0	0~999	10	0~999
2	1000~1999	0	1000~1999	10	1000~1999

【図7】

図7

記憶装置番号	記憶装置内アドレス	使用/未使用
0	0~999	使用
0	1000~1999	使用
0	2000~2999	未使用
0	3000~3999	未使用

【圖8】



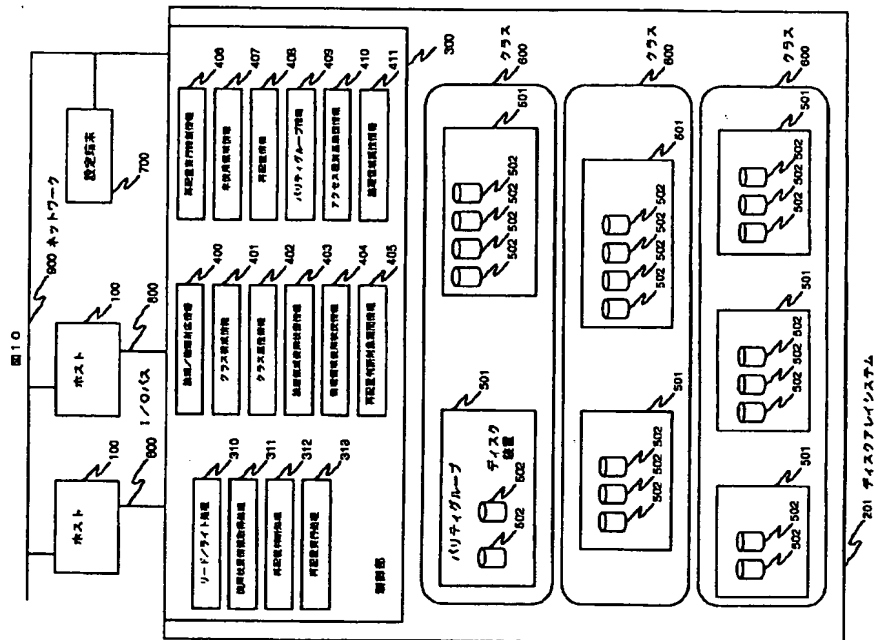
【圖 11】

投票アドレス	投票アドレス				
	パリティグループ 番号	データ		冗余データ	記憶装置内 アドレス
		記憶装置 番号	記憶装置内 アドレス		
0-999	100	0	0-999	20	0-999
1000-1999	100	0	1000-1999	20	1000-1999
2000-2999	101	1	0-999	41	0-999
3000-3999	101	1	1000-1999	41	1000-1999

【图 1-2】

クラス番号	パリティグループ数	パリティグループ番号
0	3	100, 110, 120
1	2	101, 111
2	4	102, 112, 122, 132

【圖10】



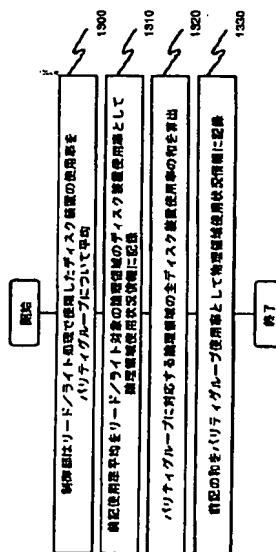
【圖 13】

13

クラス番号	使用車上座席 (%)	クラス別総席位	所配乗客数上座席 (%)	座席
0	60	1	70	—
1	70	2	80	座席
2	80	3	80	—

【图14】

14



【图 15】

15

日時	加配アドレス	テキスト郵便利用率 (%)
1999年8月1日 8時0分	0～999	18
	1000～1999	32
1999年8月1日 8時15分	0～999	20
	1000～1999	30
1999年8月1日 8時30分	0～999	22
	1000～1999	28

[圖 16]

818

日時	バリタイグループ番号	使用率(%)
1999年8月11日 8時0分	100	68
	101	52
1999年8月11日 8時15分	100	70
	101	50
1999年8月11日 8時30分	100	72
	101	48

[18]

圖 1 8

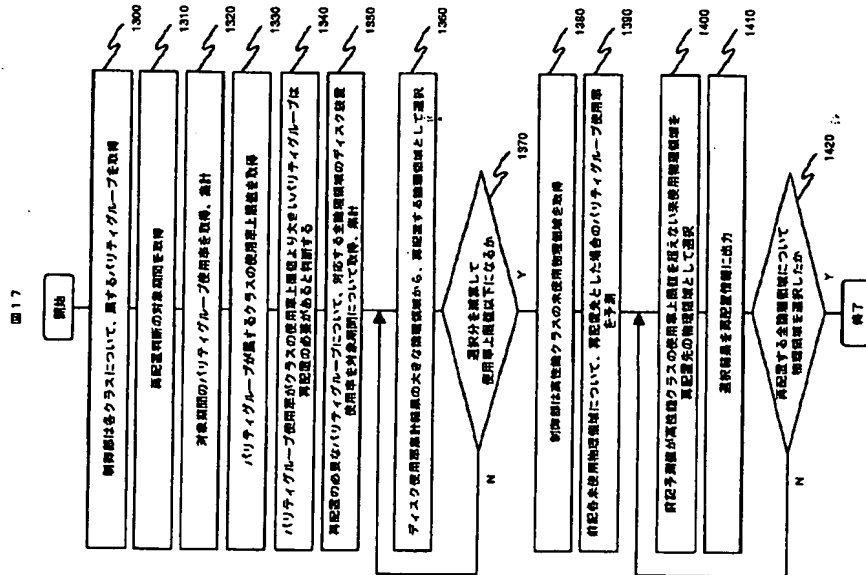
バリチグループ番号	RAID構成	ディスク数量/性能	固定
100	RAID5 3D1P	110	—
101	RAID1 1D1P	100	固定
102	RAID5 6D1P	95	—

【图22】

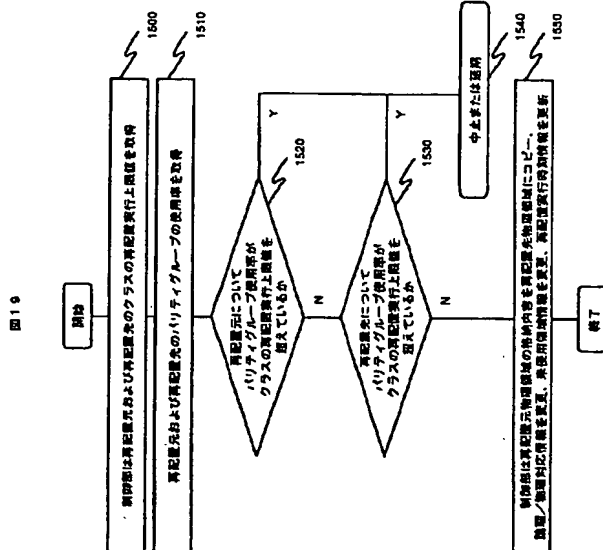
22

クラス番号	使用率上限度 (%)	クラス別 評価額立	算定 費件上限度 (%)	固定	アクセス場所
0	60	1	70	-	-
1	70	2	80	-	-
2	80	3	90	-	シーケンシャル

【図17】



【図19】

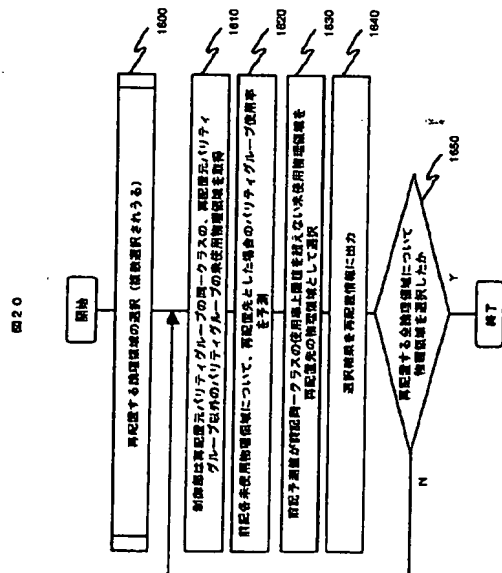


【図23】

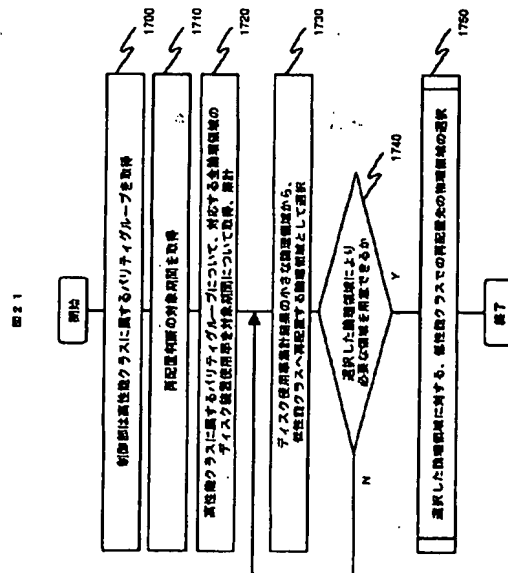
図23

日時	検索アドレス	ディスク容量 使用率 (%)	シーケンシャル アクセス率 (%)	ランダム アクセス率 (%)
1999年8月11日 8時0分	0~999	18	78	22
	1000~1999	32	82	49
1999年8月11日 8時15分	0~999	20	80	20
	1000~1999	30	80	50
1999年8月11日 8時30分	0~999	22	82	18
	1000~1999	28	48	52

【20】



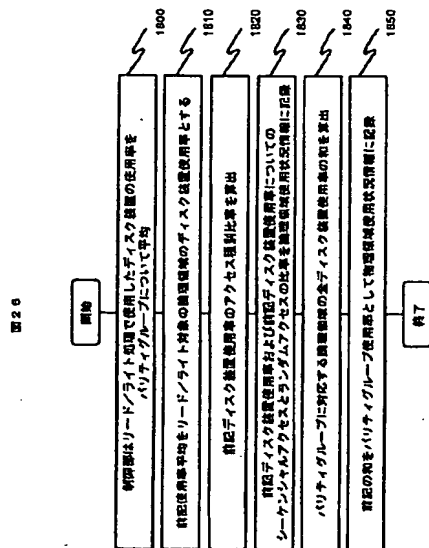
【图21】



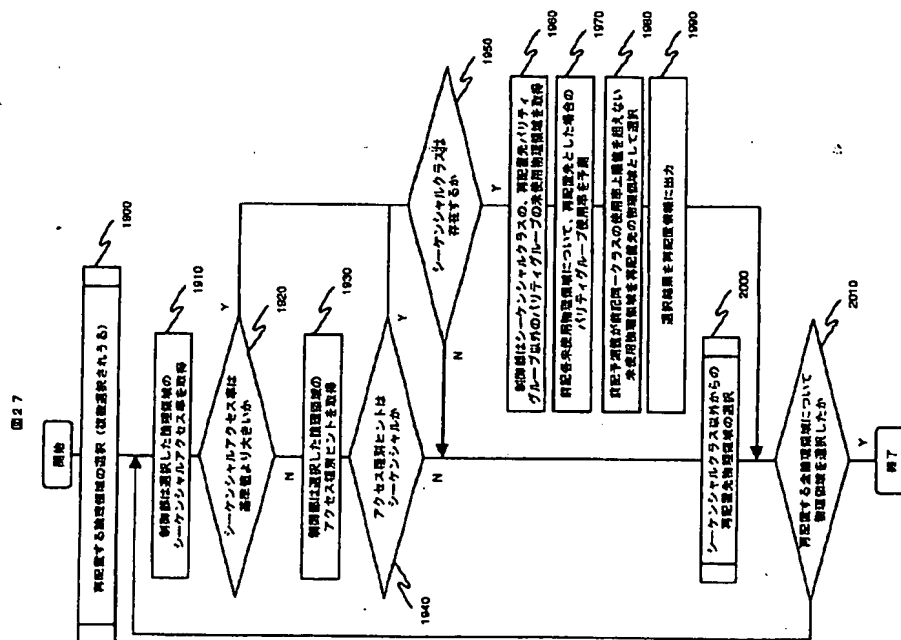
【图25】

調査アドレス	アクセス回数に占める割合	調査
0～999	—	—
1000～1999	—	—
2000～2999	シーケンシャル	—
3000～3999	—	調査

[圖26]



【图27】



フロントページの続き

(72) 發明者 山神 寔司

山仲 憲司
株 式 会 社 日 立 製 作 所 シ ス テ ム 開 発 研 究 所 内
地 番 1099 寺 王 区 麻 生 市 岐 川 県 奈 神
株

(72) 堯明者 荒井 弘治

株式会社日立製作所ストレージシステム事業部内
神奈川県小田原市国府津2880番地